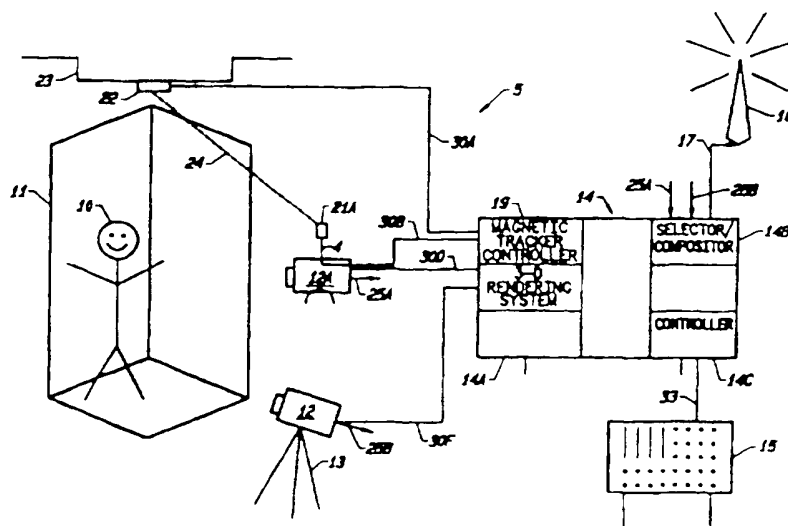




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G06T 15/00		A1	(11) International Publication Number: WO 96/32697
			(43) International Publication Date: 17 October 1996 (17.10.96)
(21) International Application Number: PCT/US96/04846		(81) Designated States: AT, AU, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, MX, NO, NZ, PL, PT, SE, SG, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 10 April 1996 (10.04.96)			
(30) Priority Data: 08/419,020 10 April 1995 (10.04.95) US			
(71) Applicant: ELECTROGIG CORPORATION [US/US]; 50 Osgood Place, San Francisco, CA 94133 (US).		Published With international search report.	
(72) Inventors: LOFTUS, James, A.; 4 Morning Sun Court, Mountain View, CA 94043 (US). REID, Ian, G.; 3316 Laguna, San Francisco, CA 94123 (US). COHEN, Steven, M.; Apartment 4, 932 Vallejo Street, San Francisco, CA 94133 (US).			
(74) Agents: WOLFELD, Warren, S. et al.; Fliesler, Dubb, Meyer and Lovejoy, Suite 400, Four Embarcadero Center, San Francisco, CA 94111-4156 (US).			

(54) Title: HAND-HELD CAMERA TRACKING FOR VIRTUAL SET VIDEO PRODUCTION SYSTEM



(57) Abstract

A virtual set video production system (5) as shown in the figure, having a hand-held camera tracking system (14) is provided. The virtual set video production system (5) includes a 3-D rendering system (14A) which generates a 3-D rendered scene and a hand-held camera (12A) which captures an image of talent. The 3-D rendering system (14A) is coupled to a magnetic tracker system (19) which provides position and orientation information of the hand-held camera (12A). The 3-D rendering system (14A) then provides a 3-D rendered scene based upon position and orientation information from the magnetic tracker system (19). A compositor (14B) then combines the video image of the talent provided by the hand-held camera (12A) with the 3-D rendered scene to produce a 3-D composite video image corresponding the hand-held camera position and orientation with respect to the 3-D rendered scene. The composite 3-D image is suitable for broadcast.

HAND-HELD CAMERA TRACKING FOR
VIRTUAL SET VIDEO PRODUCTION SYSTEM

Inventors:

5

10

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is related to U.S. Patent Application Serial No. _____, entitled INTEGRATED
15 VIRTUAL SET PRODUCTION FACILITY WITH USER INTERFACE AND CENTRAL FAULT TOLERANT CONTROLLER, by inventor James A. Loftus (Attorney Docket No. ELGG2010WSW/KJD), filed concurrently herewith and assigned to the assignee of the present application. The related
20 application is incorporated herein by reference in its entirety.

BACKGROUND

1. Field of the Invention

25 This invention relates to an apparatus and method for generating a three-dimensional ("3-D") image. In particular, this invention relates to generating a 3-D image in virtual set video production.

30 2. Description of the Related Art

A virtual set video production system refers to a combination of components used to combine or composite an image of a "real-world" object and a generated or "virtual" image. An example of a real-world object is
35 a human being, animal, inanimate object, or a combination thereof. An example of a generated image is a computer-generated 3-D rendered virtual "scene."

The rendered scene may include a plurality of computer-generated rendered images or an image stored in a storage device.

5 The "weatherman" segment which is broadcast during the evening news is typically generated by a virtual set video production system. A stationary camera on a tripod captures the image of a weatherman positioned in front of a "blue screen." Other virtual set video production components generate the virtual scene or background image which may include a map or
10 3-D rendered graphical scene.

Another example of a virtual set video production system includes a cinematic special effect application. In producing films, miniature models,
15 such as spacecraft models, may be positioned in front of a blue screen and filmed by a camera on a special track or motion control track. This specialized track must be installed and positioned. Typically, the tracks are complex and expensive. Further, the
20 special track limits or restricts the positioning and orientation of the camera because of the predefined track configuration. Eventually, the image of a real world miniature spacecraft model is combined with a generated virtual star-filled or other scene image to
25 create a realistic image of a spacecraft travelling through outer space.

As seen above, previous virtual set video production system cameras were stationary or required to be positioned and oriented on a specialized motion control track. A stationary, or restricted, camera
30 position was necessary in order to combine the real world images with the virtual images.

However, this stationary, or restricted, camera position significantly limits the composite image

which may be produced. The composite image of a real world object and virtual image is limited by the specific camera position and orientation required by a tripod or a motion control track. Also, the virtual
5 image in these composite images does not change as the camera is moved. Finally, the complex and expensive camera motion control track must be installed and maintained.

Therefore, a virtual set video production system,
10 which combines real world images, captured from a positionable camera, with virtual images to produce a composite image without the use of a tripod or specialized motion control track, is desirable.

SUMMARY OF THE INVENTION

15 According to the invention, roughly described, a method is provided which allows for generating an image by capturing a first image of an object from a positionable viewpoint. A second image is generated and combined with the first image to obtain a
20 composite image based on the object position relative to the positionable viewpoint.

In another aspect of the invention, the second image is a computer-generated 3-D rendered image and the step of capturing a first image of an object from
25 a positionable viewpoint includes obtaining a first image using a positionable sensor. The positionable sensor is moveable in a X, Y or Z direction and may be oriented about the X, Y or Z axis.

In another aspect of the invention, the
30 positionable sensor is a magnetic sensor.

In another aspect of the invention, the positionable sensor is an acoustic sensor.

In another aspect of the invention, the positionable sensor is an optical sensor.

In another aspect of the invention, the positionable sensor is a mechanical sensor.

In yet another aspect of the present invention, a method generates a 3-D image suitable for broadcast by capturing a first image of an object using a hand-held camera positioned at a distance from the object. A 3-D scene image is generated and combined with the first image to obtain a 3-D image suitable for broadcast.

According to still yet another aspect of the present invention, an apparatus for generating a 3-D image suitable for broadcast is provided. A camera which obtains an image of an object is coupled to means for determining a position and an orientation of the camera. Means for generating a 3-D scene image is based upon the position and orientation of the camera. Finally, means for compositing the image of the object and the 3-D scene image to obtain the composite 3-D image is coupled 1) to the camera and 2) the means for generating.

In another aspect of the invention, the means for determining includes a magnetic tracking system, including a magnetic transmitter and a magnetic receiver coupled to the camera. The means for generating includes a computer rendering system and the means for compositing includes a compositor or a keyer.

In another aspect of the invention, a virtual set video production system generates a 3-D image suitable for broadcast. A camera obtains an image of an individual. A magnetic tracking system is coupled to the camera and generates a camera position and camera orientation signal. A computer is coupled to the tracking system which generates a 3-D scene image responsive to the camera position and camera

orientation signal. A compositor which is coupled to the computer and the camera then combines the image of the individual with the 3-D scene image to obtain a 3-D image suitable for broadcast.

5

BRIEF DESCRIPTION OF THE DRAWINGS

This invention will be described with respect to the particular embodiments thereof, and reference will be made to the drawings, in which:

10 Fig. 1 illustrates a virtual set video production system according to the present invention;

Fig. 2 is a simplified block diagram of the virtual set video production system shown in Fig. 1;

15 Fig. 3 is a simplified block diagram of the memory head camera tripod system shown in Fig. 2;

Fig. 4 illustrates a magnetic tracker system according to the present invention shown in Fig. 1;

20 Fig. 5 illustrates an interface between a 3-D rendering system and a hand-held camera according to the present invention;

Fig. 6 illustrates a magnetic tracker system, including a transmitter and receiver, reference frame shown in Fig. 4;

25 Fig. 7 illustrates an Inventor SoPerspective Camera Node according to the present invention.

DETAILED DESCRIPTION

A. VIRTUAL SET VIDEO PRODUCTION OVERVIEW

30 Fig. 1 illustrates a virtual set video production system 5 according to the present invention. Talent 10 is positioned in front of screen 11. Talent 10 can be, for example, a human being, animal, inanimate object, or a combination thereof. There may be more than one Talent 10 positioned in front of screen 11.

The image of talent 10 will ultimately be combined or composited with a computer generated 3-D rendered image or scene which is described in detail below.

5 The image of talent 10 is captured by tripod camera 12 on tripod 13. Using tripod camera 12 on tripod 13, virtual set video production system 5 combines an image of talent 10 and a computer-generated 3-D rendered image based upon a known position of tripod camera 12. The image of talent 10
10 is transferred on bus 25(b) to selector/composer 14(b) in rack 14. Information regarding camera orientation and real-world image parameters, such as focus and zoom, are transferred on bus 30(f) to 3-D rendering system 14(a).

15 The image of talent 10 may also be captured by hand-held camera 12(a). Hand-held camera 12(a) is a hand-held camera positionable in any X, Y or Z position and in any X, Y or Z axis orientation. Hand-held camera 12(a) is typically held and controlled by
20 production personnel.

 The position and orientation of hand-held camera 12(a) is determined by a magnetic tracker system 7 comprising magnetic tracker receiver 21(a), magnetic tracker transmitter 22 and magnetic tracker controller
25 19. Magnetic transmitter 22 is positioned above talent 10 by boom 23. Receiver 21(a) is positioned approximately 1 foot above hand-held camera 12(a) by rod 4 so that reception of magnetic signals 24 are not distorted by hand-held camera 12(a) internal
30 electronics. Both receiver 21(a) and transmitter 22 are connected by bus 30(b) and 30(a), respectively, to magnetic tracker controller 19. Magnetic signals 24 are transmitted from magnetic transmitter 22 to the surrounding area in response to commands from magnetic

tracker controller 19. Hand-held camera 12(a) includes two rotary encoders for transferring image parameters, such as zoom and focus, on bus 30(d) to 3-D rendering system 14(a).

5 Rack 14 includes, among other electronic devices, 3-D rendering system 14(a), magnetic tracker controller 19, selector/composer 14(b) and controller 14(c). 3-D rendering system 14(a), along with selector/composer 14(b) and controller 14(c),
10 generates a composite 3-D rendered program feed, which includes the image of talent 10 and a 3-D rendered image generated by 3-D rendering system 14(a) on line 17. Line 17 is connected to broadcast equipment 16 which broadcasts the composite 3-D rendered video
15 image program feed.

Console 15 is also connected to controller 14(c) in rack 14 by bus 33. Console 15 allows for production personnel to control the generation of a composite program feed image, along with controlling
20 various aspects of the virtual set video production.

B. VIRTUAL SET COMPONENTS

1. 3-D Rendering System

Fig. 2 is a simplified block diagram of a virtual video set production system 5 depicted in Fig. 1. An example of a 3-D rendering system 14(a) is an
25 Onyx Workstation supplied by Silicon Graphics, Inc. ("SGI") located at 2011 N. Shoreline Boulevard, Mountain View, California 94043. A typical 3-D rendering system 14(a) is configured with 1 gigabyte
30 of Random Access Memory ("RAM"), 4 gigabytes of disk memory, 4 RM-5 raster managers, MIPS R4400 Central Processing Units ("CPUs"), a video interface module and 10 serial port sockets.

3-D rendering system 14(a) communicates with components in a virtual video set production system to obtain real-world position, real-world orientation, and image parameter information from hand-held camera 12(a) and tripod camera 12. The components include magnetic tracker system 7 (or specifically controller 19), tripod camera 12 (or specifically memory head control box 41) and hand-held camera 12(a) (or specifically, hand-held camera controller 60 providing focus and zoom information). The three components are coupled to 3-D rendering system 14(a) by bus 30(g), 30(d) and 30(f), respectively. In an embodiment, these buses are asynchronus RS-232 serial data communication buses. In an embodiment, the memory head used for tripod camera 12 is an Ultimatte Memory Head, supplied by ULTIMATTE Corp. located at 20554 Plummer Street, Chatsworth, CA 91311.

3-D rendering system 14(a) includes a software graphic engine which includes a process that updates and displays rendered 3-D images. 3-D rendering system 14(a) also includes shared memory where information from buses 30(g), 30(d) and 30(f) input data. The graphics engine has access to the shared memory to obtain data output from magnetic tracker system 7, hand-held camera 12(a) and tripod camera 12. It is essential that the graphical frame speeds are greater than 30 frames per second (National Television System Committee (NTSC) video frame speed) so that graphics rendering does not lag behind live video in the composite 3-D image.

2. Controller/Console

3-D rendering system 14(a) is connected to controller 14(c) by bus 32. The controller 14(c) is then connected to console 15 by bus 33. Controller

14(c) is also connected to selector/composer 14(b) by interface 38. A human operator will execute, configure, and modify the virtual set video production system using console 15. The human operator's inputs are transmitted on bus 33 to controller 14(c). In an embodiment, the human operator's inputs on console 15 are communicated to controller 14(c) and eventually to 3-D rendering system 14(a) on an asynchronus RS-232 serial data communication bus 32 and/or to selector/composer 14(b) on a general purpose interface 38. The list of human operator-initiated inputs include camera switching, scene setting and running animations.

One purpose of controller 14(c) is to insure a high level of reliability and performance in the system as demanded by the broadcast industry. One requirement of controller 14(c) is to insure that the program feed is not lost during system operation. A set of communication protocols are defined between controller 14(c) and 3-D rendering system 14(a). In the event of a hardware or software failure in 3-D rendering system 14(a) that results in the loss of a live program feed, 3-D rendering system 14(a) must configure to a nonrendering mode of operation. When this occurs, an auto reboot procedure will be conducted on 3-D rendering system 14(a) and the graphic engine will be restarted with the latest controller 14(c) configuration as quickly as possible. Details of the interaction between controller 14(c) selector/composer 14(b) and 3-D rendering system 14(a) with regard to enhanced reliability and performance can be found in the above-identified pending patent application entitled INTEGRATED VIRTUAL

SET PRODUCTION FACILITY WITH USER INTERFACE AND
CENTRAL FAULT TOLERANT CONTROLLER.

Virtual set video production asset management
such as switching cameras, switching scenes and
5 generating animation in the images are controlled by
console 15. Console 15 contains a number of momentary
contact pushbutton switches, each with a switch having
an associated light. This light will be lit when the
switch is activated.

10 In the present embodiment, console 15 includes a
bank of eight momentary contact lighted pushbuttons,
each corresponding to a different camera. In the
system of Fig. 1, only two of the camera pushbuttons
are used. To the right of the camera pushbuttons are
15 four rows of eight lighted pushbuttons each. Each of
the pushbuttons in a first row correspond to a
respective "virtual set", and each of the pushbuttons
in a second or third row correspond to different
animations which can be produced by 3-D rendering
20 system 14(a). The buttons in the last row are
available for future expansion. Thus, console 15
generates camera, "virtual set" and animation requests
and commands on bus 33 to controller 14(c).

a. Camera Change Requests

25 When an input on bus 33, generated by a camera
switch, is sensed by controller 14(c), the CAMERA N
command is sent to 3-D rendering system 14(a) via bus
32. The command to 3-D rendering system 14(a)
identifies which camera is being used. In the present
30 embodiment, 3-D rendering system 14(a) is told whether
tripod camera 12 or hand-held camera 12(a) is
currently being used.

After the CAMERA N code is sent, a GPI pulse will
be output on interface 38 to selector/composer

14(b). The sequence of events to produce this pulse is as follows:

Send the CAMERA N code to 3-D rendering system 14(a) via bus 32.

5 Set a "one shot" software timer to expire after the interval CurrGPIDelay. This provides a delay before GPI pulse is output on interface 38. This delay is user programmable. The value of CurrGPIDelay can be set in the range of 0 to 2200 ms. Current
10 default is 55 ms.

When the timer expires, GPI 1 On (for Tripod camera 12) or GPI 2 On (for hand-held camera 12(a)) is asserted.

A delay loop is entered next. In this loop, an
15 integer is incremented. The loop is executed GPI_PULSE_COUNT times. (For example, set at 500.)

Upon exit from the loop, GPI 1 Off (for Tripod camera 12) or GPI 2 Off (for hand-held camera 12(a)) is asserted.

20 In an embodiment, the GPI pulse output described above is generated on interface 38 to selector/compositer 14(b). In an embodiment, the GPI pulse is output to a Grass Valley Model 1000 video selector. The Grass Valley Model 1000 video selector
25 is supplied by the Grass Valley Group, Inc., Post Office Box 1114, Grass Valley, California, 94945 USA. A document describing the Grass Valley Model 1000 video selector is entitled "Grass Valley Group, Designing Digital Systems, TP3488-00, Issue B,"
30 February 1994 which is incorporated by reference herein. The GPI pulse causes either an image on line 25(b) from tripod camera 12 or the image on line 25 from hand-held camera 12(a) to be input to the compositer.

b. Set Change Commands

Controller 14(c) also receives set or 3-D image scene change commands from console 15 on bus 33. When an input on bus 33 is sensed on controller 14(c), a SET xx command is sent to 3-D rendering system 14(a) via bus 32. These set commands are used to change 3-D image scenes or sets generated by 3-D rendering system 14(a). Currently, SET 01 to SET 08 commands are implemented.

10 After the SET xx command is sent, a GPI On is asserted on bus 33 to console 15 which will cause a light associated with the Set switch to be lit.

In an embodiment, there is only one active Set at a time or only one 3-D rendered scene being generated. Accordingly, the light on the previously used Set switch will be turned off when a new Set is selected.

15 3-D rendering system 14(a) sends a SET COMMAND ACK to controller 14(c) on bus 32 when it has completed processing or generating the Set change. Animation Asset switches will be locked out (ignored) while waiting for a SET COMMAND ACK.

c. Animation Commands

Production personnel can also generate animation in a set by inputs on console 15. There may be various forms of animation effects such as 1) rotating a 3-D image object, 2) translating a 3-D object, or 3) using texture maps. Further, in an embodiment, the types of animation may be self-terminating or not requiring an ACK signal from 3-D rendering system 14(a), or alternatively, requiring an ACK signal from 3-D rendering system 14(a).

30 Type A Animations are self terminating and therefore do not require an ACK from 3-D rendering system 14(a).

As above, when an input from Type A Animation is sensed on bus 33 from console 15, an Animation Axx command is sent to 3-D rendering system 14(a) via bus 32. Currently, ANIMATION A01 to ANIMATION A08
5 commands are implemented.

After an ANIMATION Axx command is sent, a GPI On is asserted on bus 33 to console 15 which will cause the light associated with the Animation Axx switch to be lit.

10 Type B Animations are not self terminating and therefore require an ACK from 3-D rendering system 14(a).

As above, when an input from Type B Animation is sensed on bus 33 from console 15, an Animation Bxx
15 command is sent to 3-D rendering system 14(a) via bus 32. Currently, ANIMATION B01 to ANIMATION B08 commands are implemented.

After an ANIMATION Bxx command is sent, a GPI On is asserted on bus 33 to console 15 which will cause
20 the light associated with the Animation Bxx switch to be lit.

3. Video Interface Module

The 3-D rendering system includes a video interface module 14(d). In an embodiment, the video
25 interface module 14(d) is a Sirius video card. The Sirius video card is supplied by SGI, located at 2011 N. Shoreline Boulevard, Mountain View, California 94043. A document describing the Sirius video card is entitled "Sirius Video Programming and Configuration
30 Guide", which is incorporated by reference herein. Video interface module 14(d) integrates broadcast-quality video with 3-D rendering system 14(a) graphics.

Video interface module 14(d) supports real time input and output of video in the full range of broadcast video formats and supports real time input and output of video at field rates of 50, or 59.94 frames per second, in the full range of video and SGI formats. The video from video interface module 14(d) can be directed to more than one output at a time, making a broad range of processing and monitoring arrangements.

Thus, video interface module 14(d) outputs a rendered 3-D image to selector/composer 14(b) on lines 31 and 39. The talent image captured from tripod camera 12 or hand-held camera 12(a) is then output on bus 25(b) and 25(a), respectively, to selector/composer 14(b).

4. Selector/Composer 14(b)

In an embodiment, selector/composer 14(b) includes selectors or switches for selecting which camera video signal and which 3-D rendered scene will be composited. In an embodiment, the composer in selector/composer 14(b) is an ULTIMATTE-7 digital compositor. The ULTIMATTE-7 is supplied by ULTIMATTE Corp., located at 20554 Plummer Street, Chatsworth, CA 91311, U.S.A. A document describing the ULTIMATTE-7 digital compositor is entitled "ULTIMATTE-7 DIGITAL 4:2:2:4 Operating Manual, Rev.: 09/01/94," which is incorporated by reference herein.

Selector/Composer 14(b) includes a digital video image compositor or Keyer component designed to produce realistic composites of two images. Its inputs and outputs comply with CCIR-601 Serial Component Digital standard. Selector/composer 14(b) can be interfaced with any video source that can provide or accept 8 or 10 bit serial digital video,

such as video cameras, video recorders, telecines, video switchers, paint and graphic systems, etc.

There are a number of factors which can effect the realism of a composite image. Some of them
5 depend on the sophistication of the compositing device, and others are determined by production techniques including art direction and lighting.

Finally, selector/composer 14(b) combines the image of talent 10 on either line 25(a) or 25(b) and a
10 3-D rendered scene on line 39 or 31 to produce a composite image or program feed on line 17 which is ultimately sent to broadcast equipment 16 shown in Fig. 1.

15 5. Tripod Camera Memory Head/Hand-Held Camera Servo Motors

a. Tripod Camera

A memory head is used to obtain position,
20 orientation and image parameters from tripod camera 12 as shown in Fig. 3. The memory head is coupled to tripod camera 12 and 3-D rendering system 14(a) by bus 30(f). Specifically, bus 30(f) is coupled to a memory head control box 41. Power supply 40 supplies power
25 to memory head control box 41. Memory head control box 41 supplies, via bus 30(f), camera tilt, pan, focus and zoom information to 3-D rendering system 14(a). Lines 41(a), 41(b) and 41(c) are coupled to tripod camera 12. Specifically, these lines are
30 coupled to motor 42 which control focus/zoom, pan and tilt of tripod camera 12. The camera also includes focus control 43 and control handle 45 which are coupled to V/C converter box 44. Production personnel operating the camera, via control handle 45, focus
35 control 43, and keypad 47 control the focus, zoom, pan and tilt of tripod camera 12 by generating control

signals from V/C converter box 44 to memory control box 41. Memory control box 41 then generates the appropriate signals to motors 42 and also generates tripod camera 12 position and orientation, as well as
5 real-world tripod camera 12 image parameters to 3-D rendering system 14(a) on bus 30(f).

b. Hand-Held Camera

Hand-held camera 12(a) does not have a memory head as with tripod camera 12 shown in Fig. 3. Hand-
10 held camera 12(a), as shown in Fig. 5, has two rotary encoders 71 and 70 attached to the shaft of stepper motors 68 and 69, respectively, which generate real-world focus and zoom information from hand-held camera 12(a) ultimately on line 30(d) to 3-D rendering system
15 14(a). In an embodiment, line 30(d) is an asynchronous RS-232 communication data bus coupled to hand-held camera controller 60. Hand-held camera controller 60 includes interface 61. In an embodiment, interface 61 is a Kuper Board interface supplied by Kuper Controls.
20 The interface 61 communicates with stepper driver 62 on line 63 and receives information, including zoom and focus information, from encoders 70 and 71 on lines 65 and 64, respectively. Stepper driver 62 then generates control information on lines 66 and 67 to
25 stepper motors 69 and 68, respectively. Unlike tripod camera 12, position and orientation information of camera 12(a) is transferred to 3-D rendering system 14(a) by magnetic tracker system 7.

6. Magnetic Tracking System

Fig. 4 illustrates an embodiment of a tracking system shown in Fig. 1. The tracking system is a magnetic tracking system 7 which includes magnetic controller 19, transmitter 22 and receiver 21(a).

3-D rendering system 14(a) is coupled to magnetic tracker controller 19 by bus 30(g). In an embodiment, bus 30(g) is an asynchronus RS-232 communication data bus. Magnetic controller 19 includes a master 52 and possibly slave 51, 53, etc. Transmitter 22 and receiver 21(a) is coupled to master 50 by bus 30(a) and bus 30(b), respectively. Other receivers, such as receiver 21(b), may be added, depending upon the number of hand-held cameras used. As described above and illustrated in Fig. 1, receiver 21(a) is positioned approximately one foot above hand-held camera 12(a) by rod 4.

An embodiment of magnetic tracker system 7 in virtual video set production system 5 is an extended range Flock of Birds™ tracking system. The Flock of Birds tracking system is supplied by Ascension Technology Corporation Post Office Box 527, Burlington, Vermont 05402. Two documents describing the Flock of Birds tracking system is entitled "The Flock of Birds Position and Orientation Measurement System, Installation and Operation Guide, Standalone and Transmitter/Multiple Receiver Configurations", January 31, 1994 and "The Extended Range Transmitter, Supplement to the Installation and Operation Guide," March 12, 1994, which are both incorporated by reference herein.

Magnetic tracker system 7 is a six degree of freedom measuring device that can be configured to simultaneously track the position and orientation of

up to twenty-nine receivers. Each receiver is capable of making from 10 to 144 measurements per second of its position and orientation when the receiver is located within +/- 8 feet of a transmitter. Magnetic
5 tracker system 7 determines position and orientation by transmitting a pulsed DC magnetic field that is simultaneously measured by all receivers. From the measured magnetic field characteristics, each receiver independently computes its position and orientation
10 and makes this information available to 3-D rendering system 14(a).

One transmitter mounted on a pedestal in the center of a room generates sufficient signal to allow
hand-held camera 12(a), equipped with one or more
15 receiver 21(a), to generate position and orientation information in a 16 foot by 16 foot room. With four transmitters in an array, hand-held camera 12(a) may generate position and orientation information in a 24 foot by 24 foot room.

20 The most critical item in installing magnetic tracking system 7 is selecting a location for placement of transmitter 22. A poor location will result in degraded measurement accuracy.

When large metal objects are near transmitter 22
25 and receiver 21(a), the metal object will affect the accuracy of the position and orientation measurements. A large metal object is considered to be near when the distance from a transmitter to a receiver is the same as the distance from a transmitter or a receiver to
30 the large metal object. Large metal objects include metal desks, bookcases, files, floor, ceiling and walls. In non-wood commercial buildings, the floor and possibly the ceiling are constructed of concrete that contains a mesh of reinforcing steel bars. Walls

might be constructed of cinder blocks or plaster board. Plaster board walls, however, usually have internal steel supports spaced every sixteen inches. Even if a wall has no metal in it, there may be a
5 large metal object directly on the other side, such as a desk. Usually, the largest source of error is due to the floor.

One way to evaluate the surrounding effects is to install a magnetic tracking system and determine if
10 the accuracy is satisfactory. The accuracy degradation may be evaluated by taping one receiver to a cardboard box or yard stick or some other method of holding the receiver at a fixed distance above the floor. Then, a receiver may be moved farther away
15 from a transmitter in the X direction and record the receiver's Z position output. If the floor is not causing a large error, then the Z position output will remain relatively constant as the receiver moves away from the transmitter.

20 The ideal location for magnetic tracking system 7 is in an all wood building or in a large room with a stage above the floor for mounting transmitter 22 and using receiver 21(a).

Because transmitter 22 is very heavy
25 (approximately 50 lbs), fragile, and subject to performance degradation by nearby metal, the method to support the transmitter must be strong and non-metallic. Thus, boom 23 positioning transmitter 22 should have only a small amount of metal in the mount
30 such as steel bolts. Supporting transmitter 22 on a steel or aluminum framework is not desirable. Wood, structural fiberglass, or laminated phenolic for mounting transmitter 22 is preferred.

Fig. 6 illustrates the X, Y, Z coordination frame for transmitter 22 which is not coupled to boom 23. When transmitter 22 is coupled to boom 23, the +Z direction is pointing away from talent 10.

5 It should be understood that while the present invention has been described using a magnetic tracking system, there are multiple other tracking systems that could likewise be used. Other types of tracking systems would include optical, acoustic and mechanical
10 tracking systems.

An example of an optical tracking system which could be used is a MultiTrax™ optical tracker supplied by Adaptive Optics Associates, Inc., U.S.A.

15 An example of an acoustic tracker which could be used is a Wybyon Autopilot acoustic tracker supplied by Wybron, Inc., 4930 List Drive, Colorado Springs, CO 80917.

An example of a mechanical tracking system could include a moveable boom coupled to hand-held camera
20 12(a). The boom could identify position and orientation information by encoders coupled to the boom which output orientation and position information of hand-held camera 12(a). Encoders could be positioned on each axis of a cantilevered boom such as
25 a Fake Space Boom, supplied by Fake Space, Redwood City, California.

Even a global positioning satellite system using a series of satellites and a receiver mounted on camera 12(a) could be utilized to determine the
30 position and orientation of hand-held camera 12(a).

C. MAGNETIC TRACKING SYSTEM/3-D RENDERING SYSTEM INTERFACE

As shown in Fig. 4, 3-D rendering system 14(a)
35 communicates with magnetic tracker system 7 by bus

30(g). In an embodiment, bus 30(g) is a single asynchronus RS-232 communication data bus. In an embodiment which uses a single hand-held camera 12(a), a stand alone configuration magnetic tracking system is used where only master 50, transmitter 22 and receiver 21(a) are implemented. Other embodiments may include multiple slaves 52, 53, etc. and other receivers 21(b), etc. coupled to other hand-held cameras or other objects which may be tracked.

Using a single RS-232 bus to communicate with all possible receivers has the advantage of requiring less 3-D rendering hardware; it also has the disadvantage of limiting the number of measurements per second that 3-D rendering system 14(a) can obtain from each receiver. In an alternate embodiment, multiple buses 30(g) may be coupled to each master 50 and each slave (for example, 52 and 53).

Master 50 and slaves 52 and 53 communicate by an internal bus 51. To enable receivers to exchange data among themselves, each receiver is assigned a unique internal bus 51 address via the back panel dip switches. In multiple master/slave/bus configurations, RS-232 commands may be sent to a master or directly to each master/slave which has an RS-232 bus coupled to 3-D rendering system 14(a).

1. A RS-232 Command Summary

3-D rendering system 14(a) communicates with magnetic tracking system 7 by sending RS-232 commands on bus 30(g) to master 50. Each RS-232 command consists of a single command byte followed by N command data bytes, where N depends upon the command.

The RS-232 command format is as follows:

22

BIT	MS BIT		BIT								LS
	Stop	7	6	5	4	3	2	1	0		
Start											
5 RS-232 Command	1	BC7	BC6	BC5	BC4	BC3	BC2	BC1	BC0	0	

where, BC7-BC0 is the 8 bit command value and the MS
 10 BIT (Stop=1) and LS BIT (Start=0) refers to the bit
 values that 3-D rendering system 14(a) RS-232 port
 automatically inserts into the serial data stream as
 it leaves.

The RS-232 command data format is as follows:

BIT	MS BIT		BIT								LS
	Stop	7	6	5	4	3	2	1	0		
Start											
20 RS-232 Data	1	BD7	BD6	BD5	BD4	BD3	BD2	BD1	BD0	0	

where, BD7-BD0 is the 8 bit data value associated with
 25 a given command.

When using a single RS-232 bus to communicate
 with master/slaves instead of a multiple RS-232 buses,
 each RS-232 command must be prefaced with a RS-232 TO
 FBB command.

30 The following summarizes the action of some of
 the RS-232 commands generated by 3-D rendering system
 14(a). A detailed description of RS-232 commands is
 found in the above-referenced Flock of Birds™
 Installation and Operation Guide.

35

<u>Command Name</u>	<u>Description</u>
ANGLES	Data record contains 3 rotation angles.
40 ANGLE ALIGN	Aligns Bird to reference direction.
CHANGE	Changes the value of a selected

	VALUE	master/slave system parameter.
	EXAMINE	Reads and examines a selected
	master/slave	system parameter.
5	VALUE	
	HEMISPHERE	Tells master/slave desired hemisphere of operation.
10	MATRIX	Data record contains 9-element rotation matrix.
	NEXT MASTER	Selects a flock unit other than at address=1 to be master.
15	NEXT TRANSMITTER	Turns on the next transmitter in the system.
20	POINT	One data record is output for each B command from the selected master/slave unit. If GROUP mode is enabled, one record is output from all running master/slave units.
25	POSITION	Data record contains X, Y, Z position of receiver.
	POSITION/ANGLES	Data record contains POSITION and ANGLES.
30	POSITION/MATRIX	Data record contains POSITION and MATRIX.
35	POSITION/QUATERNION	Data record contains QUATERNIONS.
	REFERENCE FRAME	Defines new measurement reference frame.
40	REPORT RATE	Number of data records/second output in STREAM mode.
	RS-232 TO FBB	Use one RS-232 interface connection to talk to all master/slave.
45	SLEEP	Data records are transmitted continuously from the selected flock unit. If GROUP mode is enabled then data records are output continuously from all running
50		flock units.

XON Resumes data transmission that has been halted with XOFF.

5 XOFF Halts data transmission from master/slave.

2. RS-232 Command Operation

3-D rendering system 14(a) obtains position and orientation information regarding hand-held camera
10 12(a) by generating RS-232 commands to magnetic tracker system 7. 3-D rendering system 14(a) generates a command describing what type of data magnetic tracker system will send when a data request is issued. The desired type of data is indicated by
15 sending one of the following data record commands: ANGLES, MATRIX, POSITION, QUATERNION, POSITION/ANGLES, POSITION/MATRIX, or POSITION/QUATERNION. These commands do not cause the master/slave to transmit data to 3-D rendering system 14(a). For 3-D rendering
20 system 14(a) to receive data, it must issue a data request. The POINT data request is used each time one data record is described or a STREAM data request is used to initiate a continuous flow of data records from a master/slave. If a reduced rate at which data
25 STREAMS from a master/slave is desirable, a REPORT RATE command may be used. All commands can be issued in any order and at any time to change the master/slave's output characteristics.

The following is a typical RS-232 command
30 sequence from 3-D rendering system 14(a) to master 50, issued after power-up and configuration, which illustrates the use of some of the RS-232 commands.

For a stand-alone configuration:

	<u>COMMAND</u>	<u>ACTION</u>
5	POSITION/ ANGLES	Output records will contain certain angles or orientation) and position of receiver 21(a).
	POINT	Receiver 21(a) outputs POSITION and ANGLES data record.
10	STREAM	POSITION and ANGLE data records start streaming from master 50 and will not stop until the mode is changed to POINT.
15	POINT	A POSITION and ANGLE data record is output and the streaming is stopped.

D. 3-D RENDERING

Graphics rendering is the process of computing a two-dimensional image (or part of an image) from 3-D geometric forms. An object is considered herein to be 3-D if its points are specified with at least three coordinates each (whether or not the object has any thickness in all three dimensions). A renderer is a tool which performs graphics rendering operations in response to calls thereto. Some renderers are exclusively software, some are exclusively hardware, and some are implemented using a combination of both (e.g. software with hardware assist or acceleration). Renderers typically render scenes into a buffer which is subsequently output to the graphical output device, but it is possible for some renderers to write their two-dimensional output directly to the output device. A graphics rendering system (or subsystem), as used herein, refers to all of the levels of processing from an application level main program loop, all the way down to a graphical output device. 3-D rendering system 14(a) includes, among other software packages, two primary software packages: GL™ and Inventor™.

In alternate embodiments, other graphic software packages could also be used.

Silicon Graphics' GL is a renderer used primarily for interactive graphics. GL is described in a document entitled "Graphics Library Programming Guide," Silicon Graphics Computer Systems, 1991, which is incorporated by reference herein. It was designed as an interface to Silicon Graphics rendering hardware. GL supports simple display lists which are essentially macros for a sequence of GL commands. The GL routines perform rendering operations by issuing commands to the SGI hardware.

Inventor is an object oriented 3-D graphics user interaction toolkit that sits on top of the GL graphics system. Inventor has an entire scene or model residing in a "scene graph". Inventor has render action objects that take a model as a parameter. The render action draws the entire model by traversing the model and calling the appropriate rendering method for each object. The usual render action is the GL rendering mode. Inventor is described in a document entitled "The Inventor Mentor", Wernecke, Addison-Wesley (1994), which is incorporated by reference herein.

Inventor provides a camera class of node with a lens that functions just as the lens of a human eye does, and it also provides additional cameras that create a 2D "snapshot" of the scene with other kinds of lenses.

A camera node generates a picture of everything after it in a scene graph. Typically, you put the camera near the top left of the scene graph, since it must precede the objects you want to view. In an embodiment, a scene graph should contain only one

active camera, and its position in space is affected by the current geometric transformation.

1. SoCamera

Camera nodes are derived from the abstract base
5 class **SoCamera**.

SoCamera has the following fields:

viewportMapping (SoSFEnum)

treatment when the camera's aspect ratio is different from the viewport's aspect ration.

10 **position** (SoSFVec3f)

location of the camera viewpoint. This location is modified by the current geometric transformation.

orientation (SoSFRotation)

15 orientation of the camera's viewing direction. This field describes how the camera is rotated with respect to the default. The default camera looks from (0.0, 0.0, 1.0) toward the origin, and the
20 up direction (0.0, 1.0, 0.0). This field, along with the current geometric transformation, specifies the orientation of the camera in world space.

aspectRatio (SoSFFloat)

25 ratio of the camera viewing width to height. The value must be greater than 0.0. A few of the predefined camera aspect ratios included in *SoCamera.h* are:

SO_ASPECT_SQUARE (1/1)

30 SO_ASPECT_VIDEO (4/3)

SO_ASPECT_HDTV (16/9)

nearDistance (SoSFFloat)

distance from the camera viewpoint to the near clipping plane.

farDistance (SoSFFloat)

distance from the camera viewpoint to the far clipping plane.

focalDistance (SoSFFloat)

5 distance from the camera viewpoint to the point of focus (used by the examiner viewer).

Fig. 6 shows the relationship between the camera position, orientation, near clipping planes, far
10 clipping planes, and aspect ratio.

When a camera node is encountered during rendering traversal, Inventor performs the following steps:

1. During a rendering action, the camera is
15 positioned in the scene (based on its specified position and orientation, which are modified by the current transformation).
2. The camera creates a view volume, based on the near and far clipping planes, the aspect ratio,
20 and the height or height angle (depending on the camera type). A view volume, also referred to as a viewing frustum, is a six-sided volume that contains the geometry to be seen. Objects outside of the view volume are clipped, or thrown
25 away.
3. The next step is to compress this 3-D view volume into a 2-D image, similar to the photographic snapshot a camera makes from a real-world scene. This 2-D "projection" is now easily mapped to a
30 2-D window on the screen.
4. Next, the rest of the scene graph is rendered using this projection created by the camera.

2. SoPerspective Camera

A camera of class **SoPerspectiveCamera** emulates the human eye: objects farther away appear smaller in size. Perspective camera projections are natural in
5 situations where you want to imitate how objects appear to a human observer.

A **SoPerspectiveCamera** node has one field in addition to those defined in **SoCamera**:

heightAngle (SoSFFloat)
10 specifies the vertical angle in radians of the camera view volume.

The view volume formed by an **SoPerspectiveCamera** node is a truncated pyramid, as shown in Fig. 6. The height angle and the aspect ratio determine the width
15 angle as follows:

$$\text{widthAngle} = \text{heightAngle} * \text{aspectRatio}$$

Thus, upon every frame, 3-D rendering system
14(a), during its main loop application program, accesses shared memory in order to obtain current
20 position and orientation information of hand-held camera 12(a). The position and orientation information is continuously being transferred to shared memory by magnetic tracker system 7 on bus
30(g). The position information is then transformed
25 from receiver 21(a) position on rod 4 to hand-held camera image plane position using an offset vector based on the position of receiver 21(a) in relation to the lens of hand-held camera 12(a).

After this transformation, the position and
30 orientation information is input to the **SoPerspective** camera position and orientation fields.

Similarly, real-world image parameters from hand-held camera 12(a), such as focus and zoom, are likewise input into **SoP rspective** camera fields.

Likewise, when using tripod camera 12, actual image parameter as well as position and orientation information from tripod camera 12 are obtained by the main program in rendering system 14(a) from shared
5 memory via bus 30(f) and memory head control box 41.

The foregoing description of the preferred embodiments of the present invention has been provided for the purposes of illustration and description. It is not intended to be exhaustive or to limit the
10 invention to the precise forms disclosed. Obviously, many modifications and variations will be apparent to practitioners skilled in the art. The embodiments were chosen and described in order to best explain the principles of the invention and its practical
15 applications, thereby enabling others skilled in the art to understand the invention for various embodiments and with the various modifications as are suited to the particular use contemplated. It is intended that the scope of the invention be defined by
20 the following claims and their equivalents.

CLAIMS

What is claimed is:

1) A method for generating an image, comprising the steps of:

capturing a first image of an object from a positionable viewpoint;

generating a second image; and,

combining the first image with the second image to obtain a composite image based on the object position relative to the positionable viewpoint.

2) The method of claim 1, wherein the second image is a computer-generated image.

3) The method of claim 2, wherein the computer-generated image is a three-dimensional image.

4) The method of claim 1, wherein the step of capturing a first image of an object from a positionable viewpoint includes obtaining the first image using a positionable sensor.

5) The method of claim 4, wherein the positionable sensor is moveable in a X, Y or Z direction.

6) The method of claim 4, wherein the positionable sensor is oriented about a X, Y or Z axis.

7) The method of claim 4 wherein the positionable sensor is a magnetic sensor.

8) The method of claim 4, wherein the positionable sensor is an acoustic sensor.

9) The method of claim 4 wherein the positionable sensor is an optical sensor.

10) The method of claim 4, wherein the positionable sensor is a mechanical sensor.

11) The method of claim 1, wherein the composite image is a video image suitable for broadcast.

12) A method for generating a three-dimensional image suitable for broadcast, comprising the steps of:
capturing a first image of an object using a hand-held camera;
generating a three-dimensional scene image;
and,

combining the first image with the three-dimensional scene image to obtain a three-dimensional image suitable for broadcast responsive to a position and an orientation of the hand-held camera.

13) The method of claim 12, wherein the hand-held camera is positionable in a X, Y or Z direction.

14) The method of claim 12, wherein the hand-held camera is oriented about a X, Y or Z axis.

15) The method of claim 12, wherein the three-dimensional scene image is computer-generated.

16) The method of claim 12, wherein the hand-held camera is coupled to a magnetic tracking system.

17) The method of claim 12, wherein the hand-held camera is coupled to an acoustic tracking system.

18) The method of claim 12, wherein the hand-held camera is coupled to an optical tracking system.

19) The method of claim 12, wherein the hand-held camera is coupled to a mechanical tracking system.

20) An apparatus for generating a three-dimensional video image suitable for broadcast, comprising:

a camera obtaining a video image of an object;

means, coupled to the camera, for determining a position and orientation of the camera;

means for generating a three-dimensional scene video image based upon the position and orientation of the camera; and,

means, coupled to 1) the camera and 2) the means for generating, for compositing the video image of the object with the three-dimensional scene video image to obtain a composite three-dimensional video image.

21) The apparatus of claim 20, wherein the means for determining includes a magnetic tracking system.

22) The apparatus of claim 21, wherein the means for determining includes a magnetic transmitter and a magnetic receiver coupled to the camera.

23) The apparatus of claim 20, wherein the means for determining includes an acoustic tracking system.

24) The apparatus of claim 20, wherein the means for determining includes an optical tracking system.

25) The apparatus of claim 20, wherein the means for determining includes a mechanical tracking system.

26) The apparatus of claim 20, wherein the means for generating includes a rendering system.

27) The apparatus of claim 26, wherein the rendering system includes a computer.

28) An apparatus of claim 20, wherein the means for compositing includes a compositor.

29) An apparatus of claim 20, wherein the means for compositing includes a keyer.

30) The apparatus of claim 20, wherein the means for determining generates the camera X, Y, Z position and X, Y or Z axis orientation.

31) A virtual set video production system generating a three-dimensional image suitable for broadcast, comprising:

a camera obtaining an image of an individual;

a tracking system, coupled to the camera, generating a camera position and camera orientation signal; and

a computer, coupled to the tracking system, generating a three-dimensional scene image responsive to the camera position and camera orientation signal; and,

a compositor, coupled to the computer and the camera, combining the image of the individual with the three-dimensional scene image to obtain a three-dimensional image suitable for broadcast.

32) The virtual set video production system of claim 31, wherein the tracking system is a magnetic tracking system.

33) The virtual set video production system of claim 31, wherein the tracking system is an acoustic tracking system.

34) The virtual set video production system of claim 31, wherein the tracking system is an optical tracking system.

35) The virtual set video production system of claim 31, wherein the tracking system includes a mechanical tracking system.

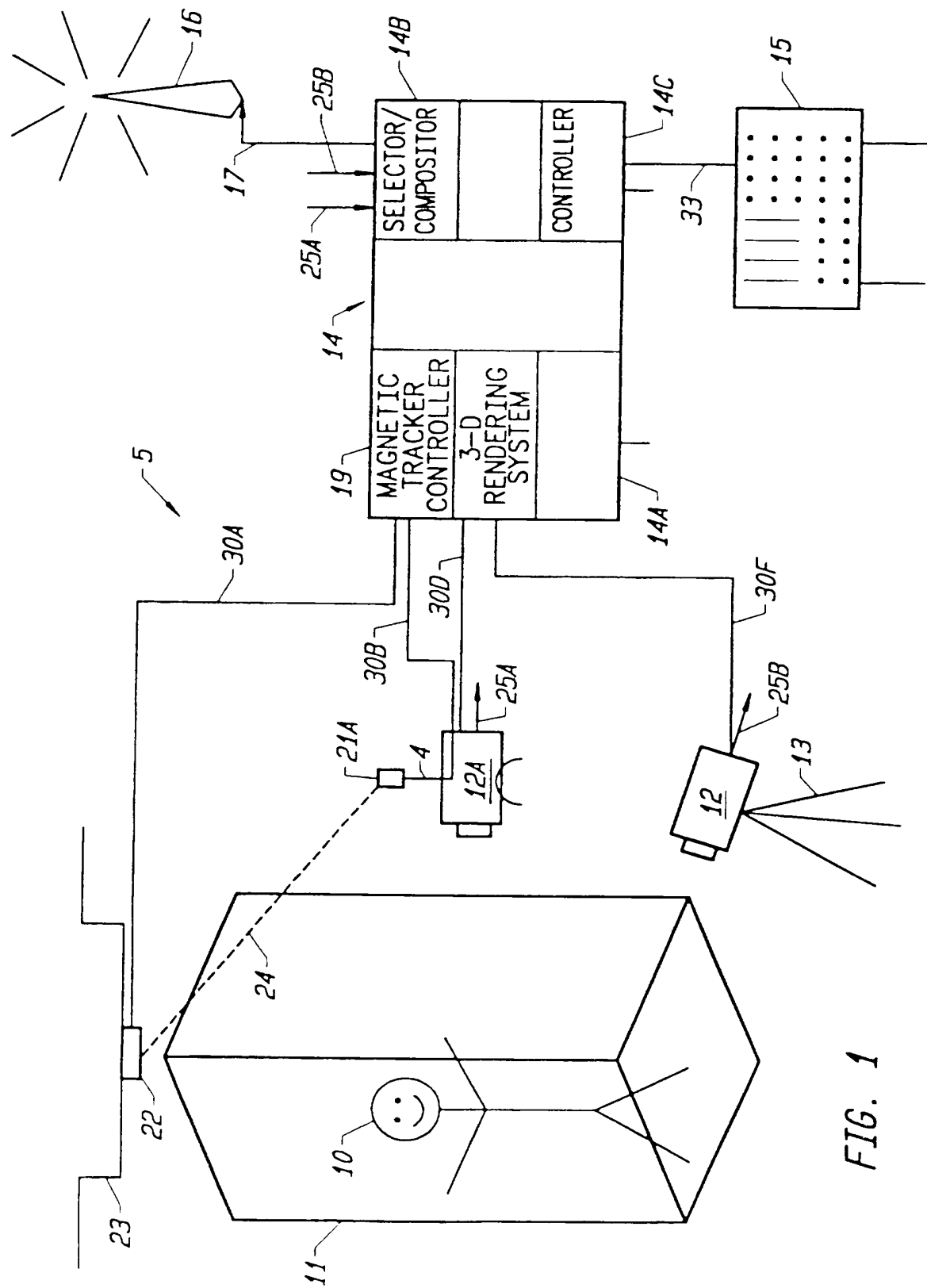


FIG. 1

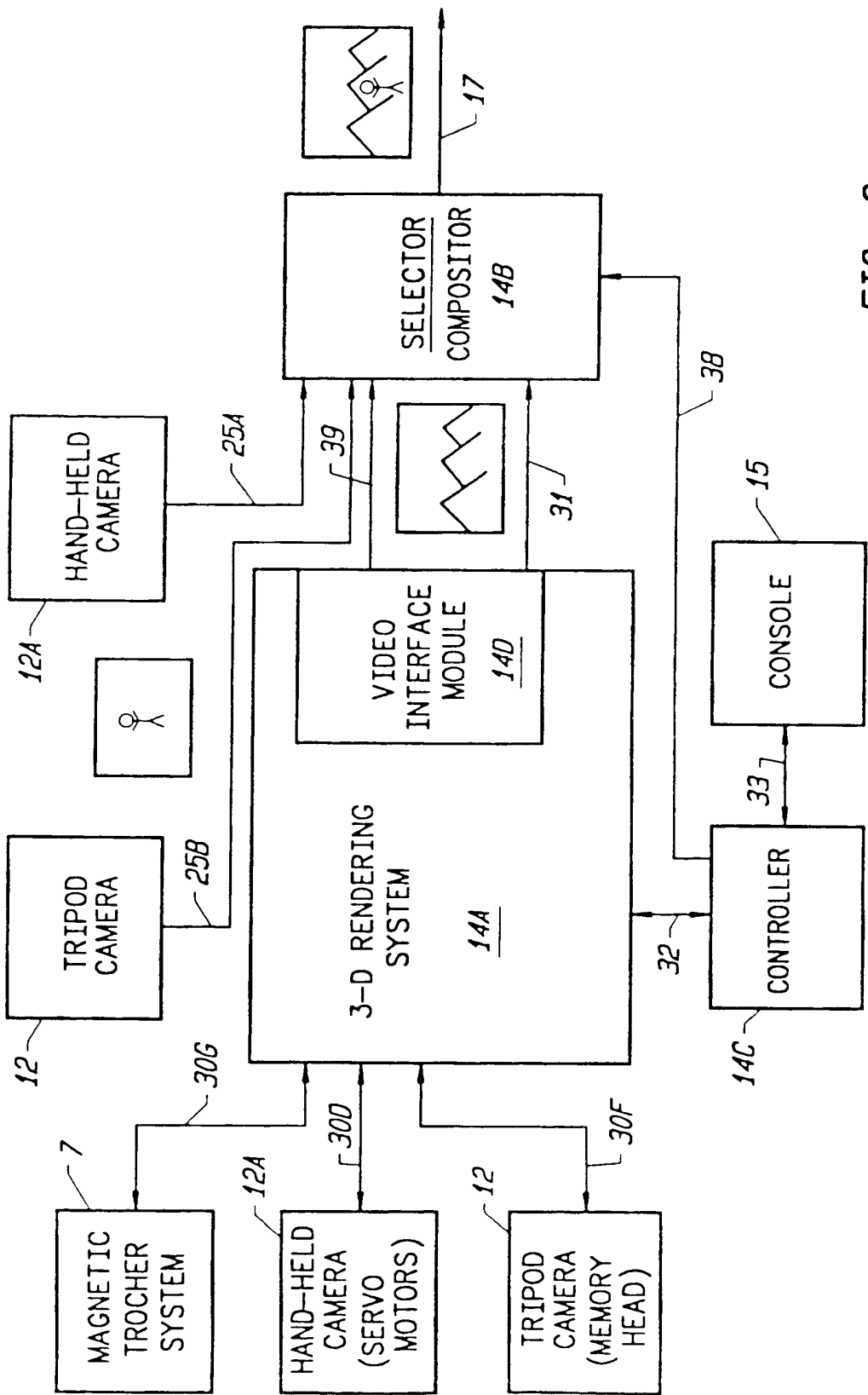


FIG. 2

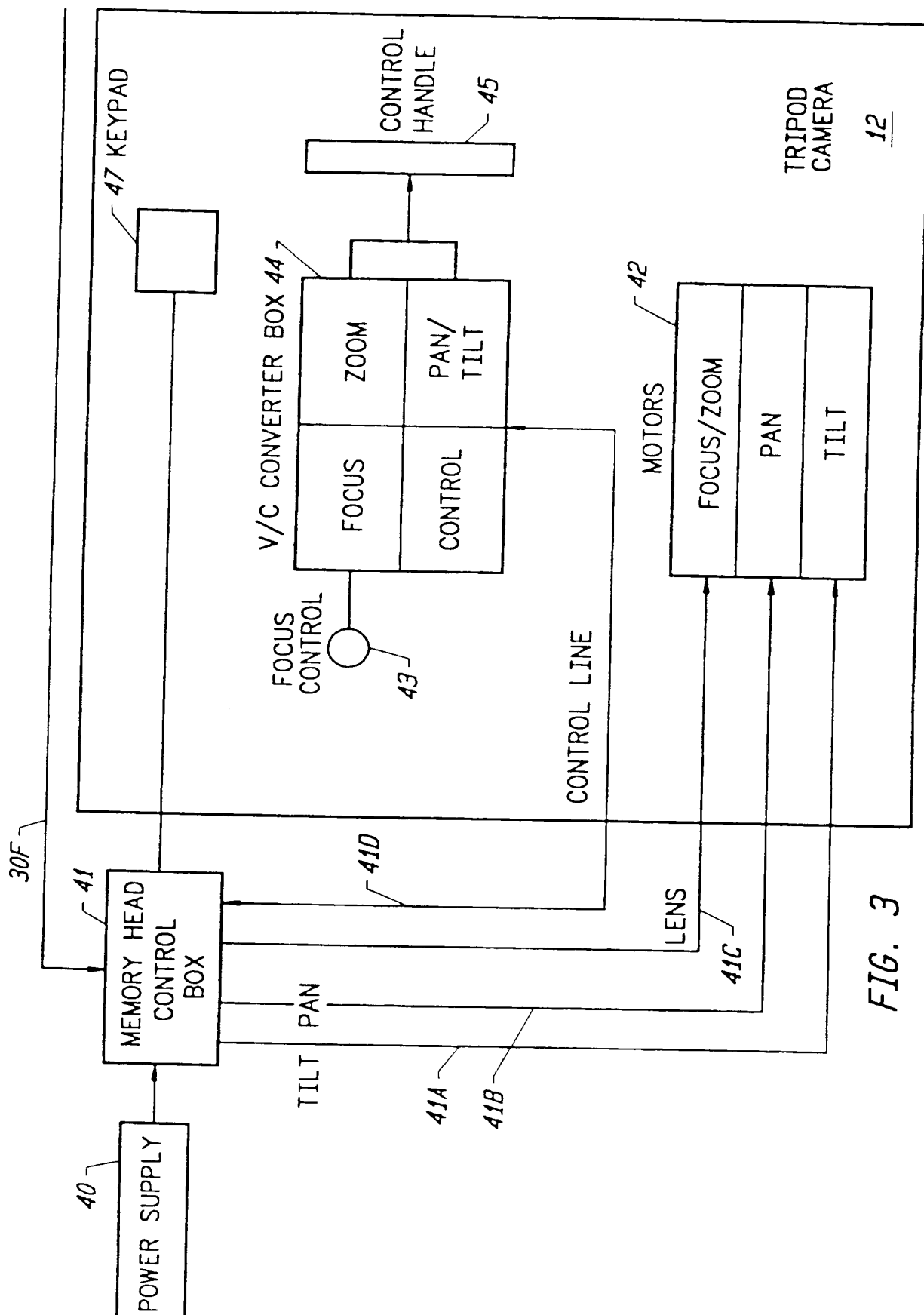


FIG. 3

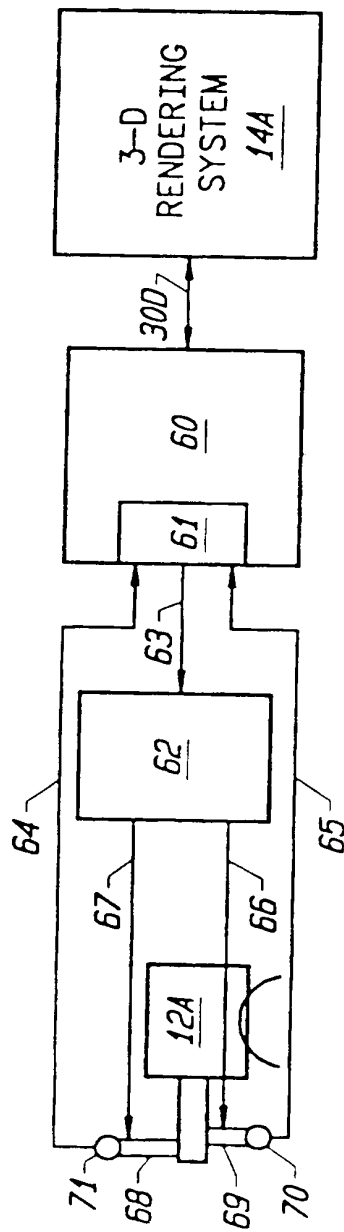


FIG. 5

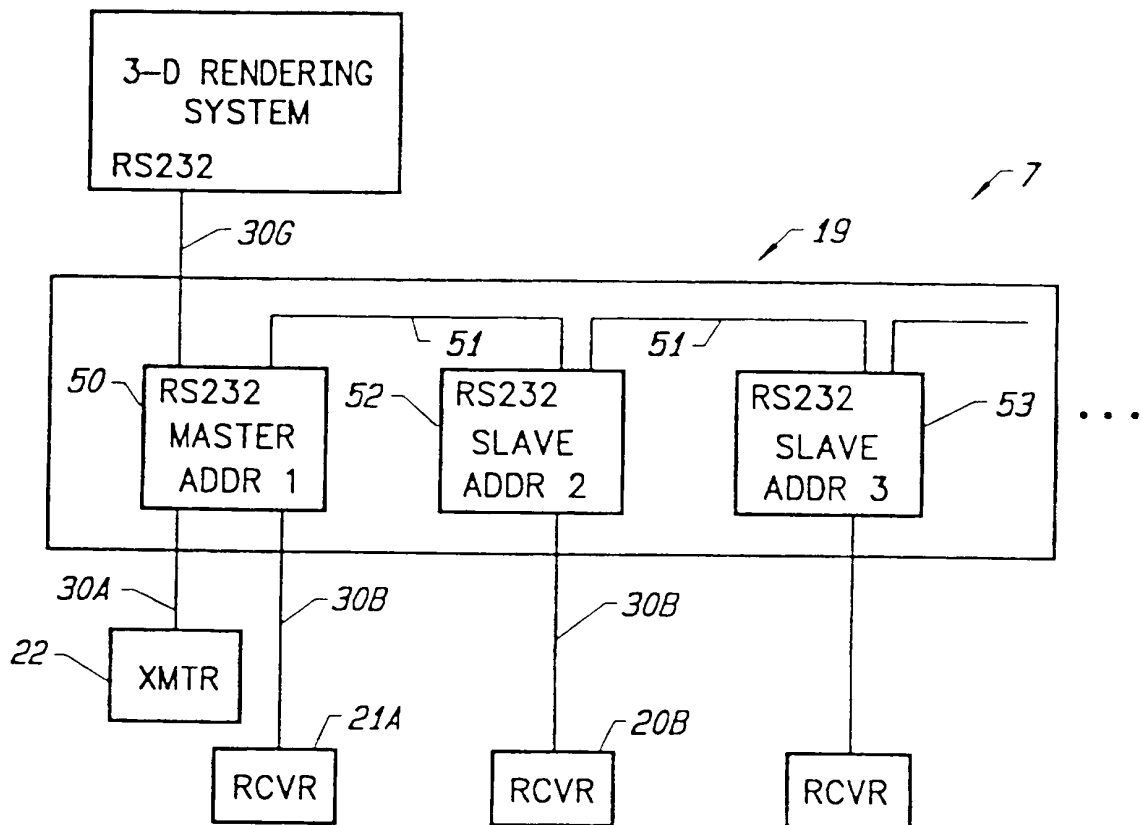


FIG. 4

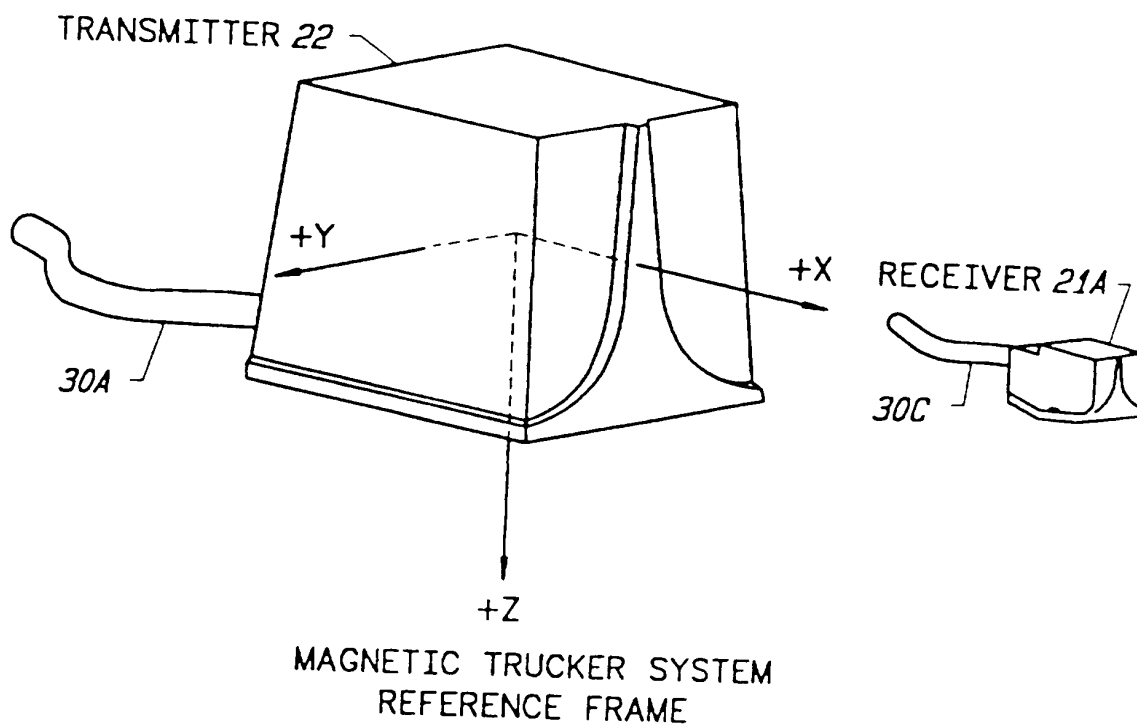


FIG. 6

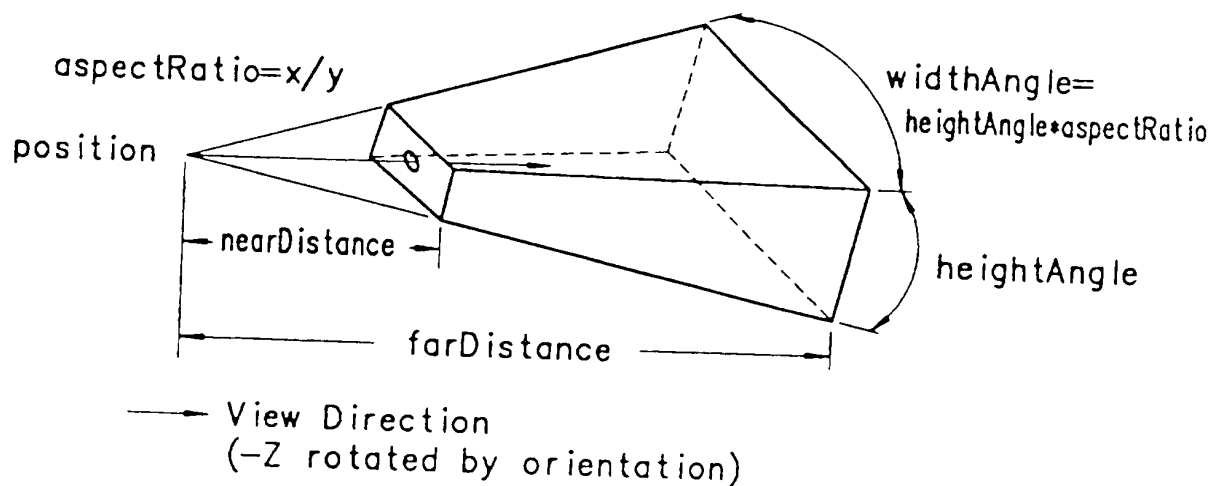


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/04846

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : G06T 15/00

US CL : 395/119

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 395/119, 395/121

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

IEEE Publications OnDisc; DIALOG online services - ELETRON, SOFTWARE, EECOMP, PCINFO and COMPSCI databases

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Please See Continuation of Second Sheet.	

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be part of particular relevance	*X*	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y*	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reasons (as specified)	*A*	document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means		
P document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

18 JUNE 1996

Date of mailing of the international search report

25 JUL 1996

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Authorized officer
RUDOLPH J. BUCHEL

Facsimile No. (703) 305.3730

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/04846

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X Y	Deering, Michael, "Explorations of Display Interfaces for Virtual Reality", Virtual Reality, 1993 International Symposium, pp. 141-147; Pages 143-145, section 5.0 Computer Augmented Reality Camera.	1-6, 12-15, 26-31 7-11, 16-25 and 31-35
Y	State, et al, "Case Study: Observing a Volume Rendered Fetus within a Pregnant Patient", Visulations, 1994 Conference pp. 364-368; page 364, second column, section 3 Near-real-time visualization system; page 365, first column, line 19, section 4.1 Calibration and Figure 1.	1-35
X Y	Azarbayejani, et al, "Recursive Estimation for CAD Model Recovery", CAD-Based Vision, 1994 Workshop, pp. 90 - 97. Page 90, column 1, Introduction and page 95, section 4 Results and Figure 3.	1-6, 12-16 and 26-31; 7-11, 17-25 and 32-35